+ Schedule

- Review project overview and goals
- Show Million Song Dataset
- Describe data mining algorithm implementation
  - Association Rules
  - Collaborative Filtering
  - Naïve Bayes
- Address testing methodology and quality
- Provide recommendation
- Questions

+ Overview

- Millions of people buy and listen to songs on the Internet every year [1]
- Music providers try to engage customers by recommending new songs to them[2]
  - Increased engagement -> song purchases = increased revenue
- Challenge: What is the best algorithm to recommend new songs to music listeners?
- Solution: Analyze three common data mining algorithms to determine the best fit

Million Song Dataset

Million Song Dataset

- 1 Million popular songs curated by LabROSA (Columbia University)[3]
- Taste Profile Subset: Real song listening transactions provided by The Echo Nest
  - Over 48 million transactions
  - Over 1 million unique listeners
  - 380,000 distinct songs
- Every user listened to at least 10 songs (avg = 48 & max = 9,600!)
- Most listened to song? Katy Perry, Firework

Taste Profile Subset

Anonymous User ID	Song ID	# Plays
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOAKIMP12A8C130995	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOAPDEY12A81C210A9	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBBMDR12A8C13253B	2
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBFNSP12AF72A0E22	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBFOVM12A58A7D494	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBNZDC12A6D4FC103	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBSUJE12A6D4F8CF5	2
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBVFZR12A6D4F8AE3	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBXALG12A8C13C108	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBXHDL12A81C204C0	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOBYHAJ12A6701BF1D	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SOCNMUH12A6D4F6E6D	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SODACBL12A8C13C273	1
b80344d063b5ccb3212f76538f3d9e43d87dca9e	SODDNQT12A6D4F5F7E	5

Transactions

Data Mining Algorithms

Association Rules (AR)

- Market-basket analysis of song play transactions
- Used HANA to generate 2-itemset ARs
  - Combined set of ARs for a song determines predictions ranked by confidence
  - Expensive due to the number of frequent items
  - Generating itemsets > 3 on HANA led to out-of-memory errors

Rule	Confidence
SOXKSAJ12A8C14588D => SOGDNKT12A8C1447DA	0.4663072776280323
SOZZWDF12A8C14144C => SOKPDCL12A8C13B5E7	0.6455223880597015
SOFLQTC12A67021CCA => SOCHRFO12AF729B18C	0.4196891191709845

- Algorithm assumptions:
  - Max Item Set = 2
  - Minimum Support = 0.00001 (~ 10 users played song)
  - Minimum Confidence = 0.20
- Generated ~1,500,000 rules

Collaborative Filtering (CF)

- Crowd based recommender algorithm which aggregates users' song selections

Song	A	B	C	D
User 1	1	1	1	0
User 2	0	1	0	1
User 3	1	1	0	1

- Example:

A New User listens to Song D, and we need to recommend songs:

- Find song D in matrix and users who listened to it: U{2, 3}
- Identify other songs users (U) listened to and rank

Recommend Song	Count of User	Rank
B	2	1
A	1	2

- Result: 1st Song Recommended B, 2nd is A, with none for C

Naïve Bayes (NB)

- Uses probabilities to classify and rank recommendations

- Example:

Calculate probabilities User listened to song A & D:

User Play History

Song	A	B	C	D
User 1	1	1	1	0
User 2	0	1	0	1
User 3	1	1	0	1

  - P(Listened to Both/Total User Just D) = 1/2 = 50%
  - P(Not listened) = (Total Users – Both)/Total Users = (3 -1)/3 = 66%
  - **Adjusted Probability = 50%/(50%+66%) = 43.1%**

- Algorithm assumptions: Songs with very few plays (<10% of plays vs. driver) removed to prevent unrealistic recommendations

Quick comparison of models

Driver Song: U2 – Endless Deep

Collaborative Filtering

title	play_count	rank
Dancing Barefoot	175	1
Hold Me_ Thrill Me	175	2
Love Comes Tumbl	164	3
Walk To The Water	144	4
A Day Without Me	84	5
Window In The Ski	82	6
I Still Haven't Foun	78	7
Bad	76	8
Vertigo	57	9
Mysterious Ways	55	10

Naïve Bayes

title	adjust_p	rank
Dancing Barefoot	22%	1
Walk To The Water	20%	2
A Day Without Me	7%	3
Love Comes Tumbl	4%	4
Hold Me_ Thrill Me	3%	5

Association

title	confidence	rank
Dancing Barefoot	45%	1
Hold Me_ Thrill Me	45%	2
Love Comes Tumbl	42%	3

Testing Methodology and Quality Assessment

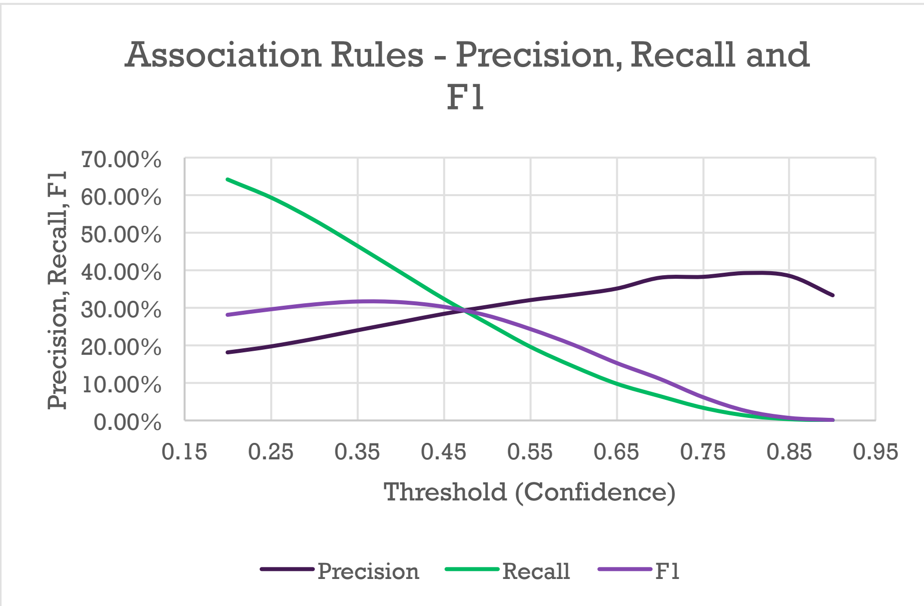
Testing Methodology

- Split the Taste Profile Subset into two datasets:
  - Training dataset containing transactions for ~1 million users
  - Test dataset containing transactions for 1 thousand users

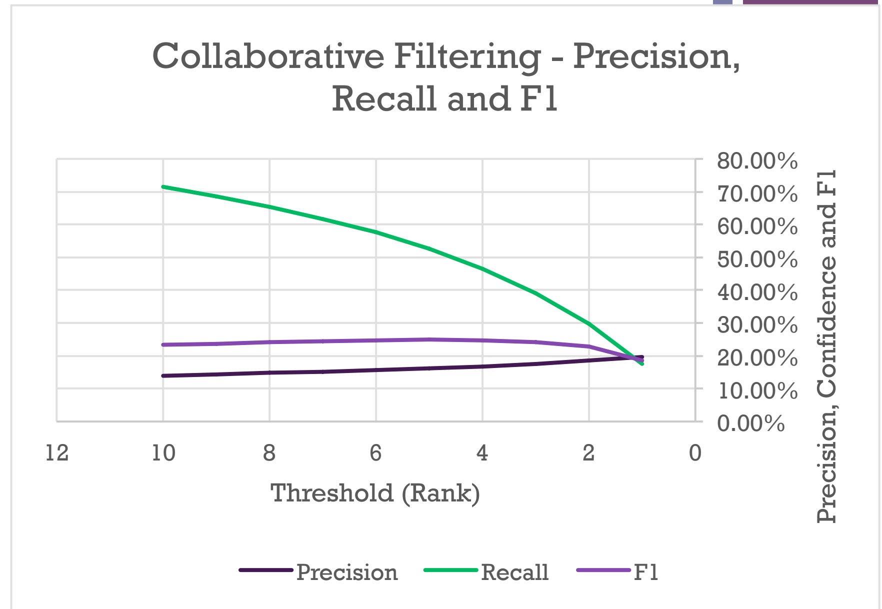
- Generate predictions using each algorithm for each user in the testing dataset

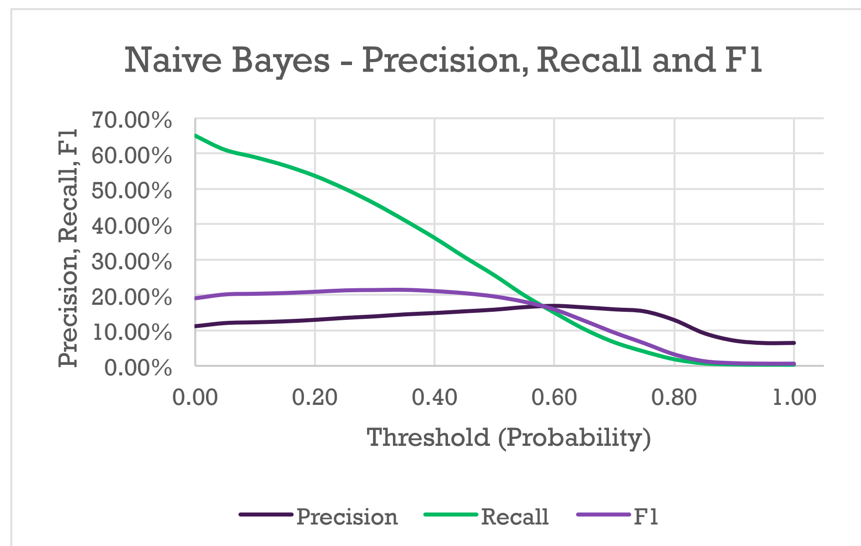
- Tested predictions:
  - Split Test Dataset into Driver Set and Target Set
  - Determined how well our Driver Set could predict our Target Set
  - Created confusion matrices at varying thresholds to quantify algorithm comparison

Testing Results



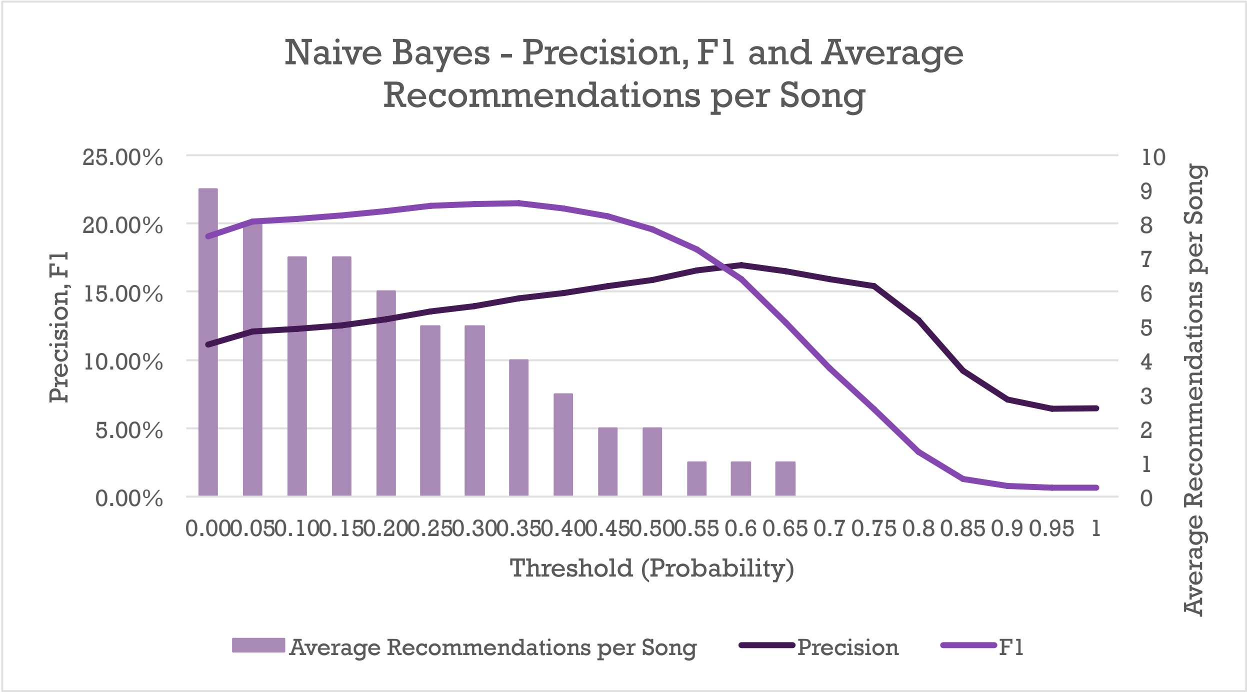
- Collaborative Filtering exhibits a relatively gradual decline in precision as the threshold is made less stringent across the test range



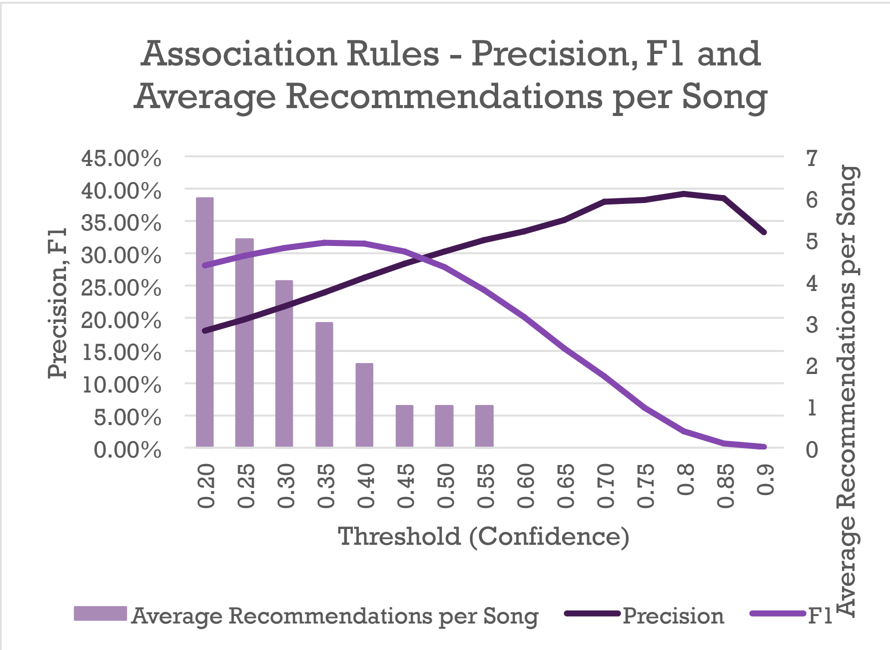


+ Testing Results

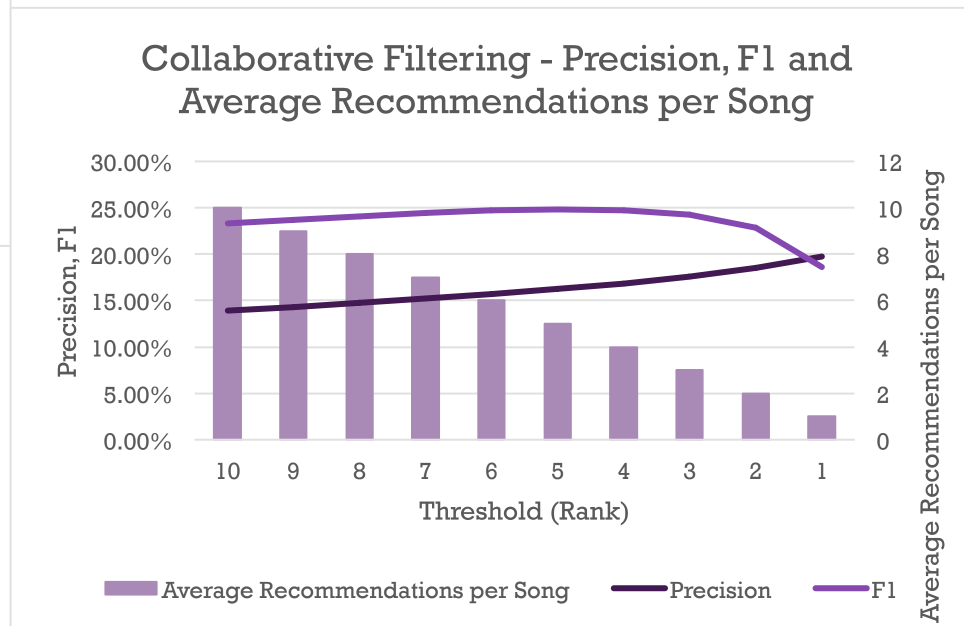
- **Business Imperatives:**
  - Keep user engaged by providing relevant recommendations → **PRECISION**
  - Provide a reasonable number of recommendations to give user options
    - Use metric directly impactful to actual business application → **AVERAGE RECOMMENDATIONS PER SONG**
    - Strike a balance between the two measures



Testing Results

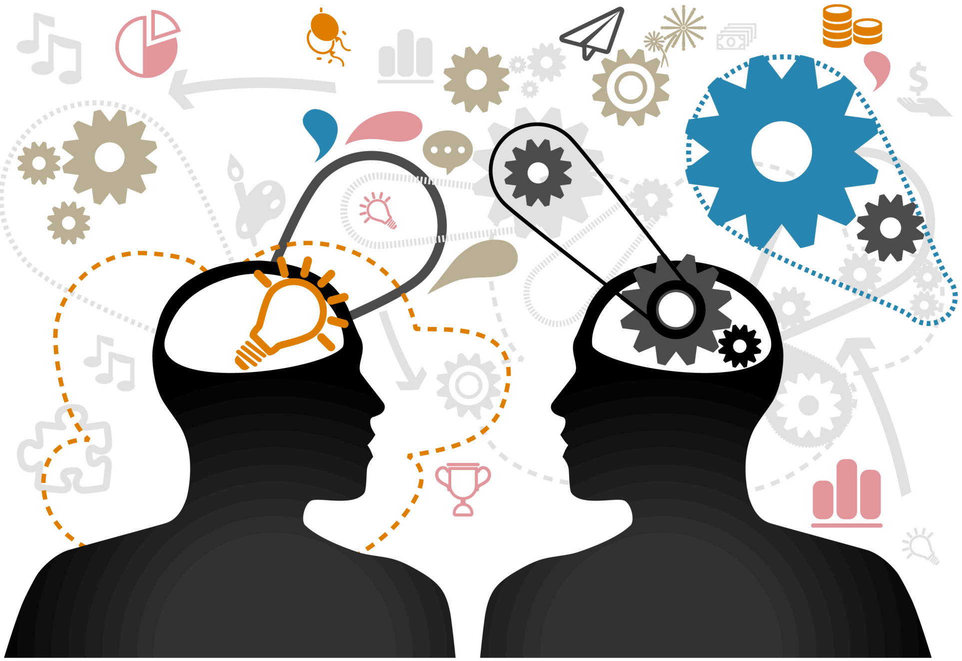


- **Collaborative Filtering**
  - Exhibiting a gradual decline in precision in the test data
  - *Possibly scalable based on the particular requirements of application*



- **Association Rules**
  - Higher precision for fewer recommendations but exhibiting a sharper decline.
  - *Seemingly better suited to formats where fewer recommendations are passable*

Recommendation and lessons learned



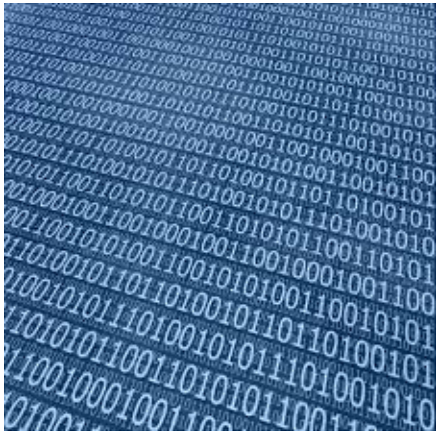
Recommendations

- Naïve Bayes
  - Precision not favorable at comparable level of recommendations
  - Dead-zones
- Collaborative Filtering
  - Sustained precision
  - Scales well for larger datasets
  - Inexpensive computationally
- Association Rules (✓ Our Recommendation)
  - Highest attained precision but sharper decline with increase in recall
  - Dead-zone of average recommendations per song

Lessons learned

- Be prepared to iterate through the entire process rather than just within components
  - Development, implementation and testing
  - High volume of short life-cycle code
  - Establish clear methodology—plan multiple iterations
  - Experimentation necessary to discover interesting trends and patterns

+



Questions & Answers

+
Citations

1. Sisario, B. (2014, September 25). U.S. Music Sales Drop 5%, as Habits Shift Online. Retrieved May 2, 2015, from http://www.nytimes.com/2014/09/26/business/media/music-sales-drop-5-as-habits-shift-online.html?_r=0

2. Kaufman, Jaime C., "A Hybrid Approach to Music Recommendation: Exploiting Collaborative Music Tags and Acoustic Features" (2014). *UNF Theses and Dissertations.* Paper 540.

3. Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. In Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011.